Sparse coded image super-resolution using K-SVD trained dictionary based on regularized orthogonal matching pursuit

Muhammad Sajjad, Irfan Mehmood and Sung Wook Baik[*]
Digital Contents Research Institute, Sejong University, Seoul, Korea

Abstract. Image super-resolution (SR) plays a vital role in medical imaging that allows a more efficient and effective diagnosis process. Usually, diagnosing is difficult and inaccurate from low-resolution (LR) and noisy images. Resolution enhancement through conventional interpolation methods strongly affects the precision of consequent processing steps, such as segmentation and registration. Therefore, we propose an efficient sparse coded image SR reconstruction technique using a trained dictionary. We apply a simple and efficient regularized version of orthogonal matching pursuit (ROMP) to seek the coefficients of sparse representation. ROMP has the transparency and greediness of OMP and the robustness of the L_1-minization that enhance the dictionary learning process to capture feature descriptors such as oriented edges and contours from complex images like brain MRIs. The sparse coding part of the K-SVD dictionary training procedure is modified by substituting OMP with ROMP. The dictionary update stage allows simultaneously updating an arbitrary number of atoms and vectors of sparse coefficients. In SR reconstruction, ROMP is used to determine the vector of sparse coefficients for the underlying patch. The recovered representations are then applied to the trained dictionary, and finally, an optimization leads to high-resolution output of high-quality. Experimental results demonstrate that the super-resolution reconstruction quality of the proposed scheme is comparatively better than other state-of-the-art schemes.

Keywords: Image super-resolution, sparse representations, reconstruction, dictionary

1. Introduction

Applications in various fields, such as medical imaging and high-resolution (HR) displays, need resolution enhancement of the images captured by low-quality visual sensors [1, 2]. In medical imaging, diverse physical, technological, and economical consideration limit the resolution of magnetic resonance (MR) images. This limitation of resolution introduces a series of bad artifacts in MR images, such as partial volume (PV) effect that degrades the performance of post processing algorithms and image analysis tasks, preventing derivation of accurate measurement and diagnosing [3]. Traditional interpolation techniques such as bicubic or B-spline are usually used to enhance the resolution of MR images, but these techniques cannot properly reconstruct the sharp luminance and complex texture of the underlying MR images and strongly affect the subsequent processing steps,

* Address for correspondence: Sung Wook Baik, Digital Contents Research Institute, Sejong University, Seoul, Korea. Tel.: +82-02-3408-3797; Fax: +82-02-3408-4339; E-mail: sbaik@sejong.ac.kr.

such as MR image segmentation and registration. Hence, any super-resolution (SR) reconstruction scheme should have the capability to coherently recover these sharp luminance and complex texture. In image SR, reconstruction constraint (SC) and sparsity prior (SP) are two important components as described by [4]. SP demands that the reconstructed HR image $A \in R^H$ must be consistent with low-resolution (LR) image $B \in R^L$ with-respect-to the image observation model. A is the HR version of the corresponding observed LR image B, where B is obtained by downsampling A as:

$$B = Q_{(d)}A + \varepsilon_{(d)}$$

(1)

where $Q_{(d)}$ is the interpolation operator for downsampling and $\varepsilon_{(d)}$ is the interpolation error or blurring effect. In SP, HR image A can be well represented as sparse linear combination of the prototype signal atoms determined from dictionary $G \in R^{n \times m}$.

$$A \approx GX \quad \text{where} \quad \min_X \sum_i \|X\|_0 \; : \; \|B - GX\|_F^2 \leq \varepsilon$$

(2)

where ε is a fixed value indicating the limit for the reconstruction error. A similar objective could be achieved alternatively by considering the problem posed in (2) as sparsity constraint

$$A \approx GX \quad \text{where} \quad \min_X \|B - GX\|_F^2 \; : \; \sum_i \|X\|_0 \leq \mathcal{K}$$

(3)

where $\|x\|_0$ is simply the number of nonzero elements of $X = \{x_1, \ldots, x_n\} \in R^{m \times n}$ and K is the threshold for maximum number of atoms. Unfortunately, solving the problem posed in Eq. (3) involves a combinatorial search, making it computationally intractable [5]. The L_1-minization, as the closest convex function to L_0-minization, is extensively used as an alternative scheme to solve the problem in Eq. (3) [6]. In other words, the two major algorithmic approaches to solve problem posed in Eq. (3) are L_1-minization and L_0-minization (matching pursuits). The problem in Eq. (3) can also be denoted as a sparse recovery problem. Yang *et al.* [4] solved the same problem using L_1-minization known in statistical literature as the Lasso. They incorporated the work of Lee *et al.* [7] for dictionary learning. Zeyde *et al.* [8] used orthogonal matching pursuit (OMP) to recover the sparse representation both in dictionary learning and SR. They applied K-SVD dictionary training procedure [9] for dictionary learning. In [10], author applied the same approach for dictionary learning, however they determined the sparse coefficients by solving the problem in Eq. (3) using L_1-minization. In our previous scheme, Sparse Representations-based Super-Resolution of Key-Frames Extracted from Frames-Sequences Generated by a Visual Sensor Network (SR-KF) [11], optimized orthogonal matching pursuit (OOMP) was applied for sparse-representation recovery in SR. For dictionary, the K-SVD dictionary training procedure was incorporated. In literature, it has been proved experimentally that L_1-minization can produce more precise solutions than matching pursuits (L_0-minization) but usually needs higher time complexity [12]. Therefore, there must be a minimization method to solve the problem posed in (3) that has advantages of both approaches (i.e., L_0-minization and L_1-minization).

In this paper, we present a sparse representation based image SR reconstruction scheme for brain MR images. First, the dictionary is trained using the modified version of the K-SVD dictionary learning procedure. The sparse coding phase of the K-SVD dictionary learning scheme is then enhanced incorporating a simple and an efficient regularized version of orthogonal matching pursuit (ROMP) [13, 14]. ROMP has the transparency and greediness of OMP and the robustness of the

L_1-minization, and this property of ROMP enriches the learning process of the K-SVD dictionary learning scheme in terms of features representation such as oriented edges and contour. In addition, the dictionary update stage allows simultaneously updating an arbitrary number of atoms and vectors of sparse coefficients. In SR reconstruction, ROMP is used to seek the vector of sparse representation coefficients for the underlying patch. The sought vector of sparse coefficient is then multiplied to the trained dictionary to reconstruct the HR output of the underlying LR patch. The final HR output of high quality is generated by solving the optimization problem with respect to trained dictionary. Experimental results validate the qualitative and quantitative significance of the proposed SR reconstruction scheme over other state of the art schemes.

The remainder of this paper is organized as follows: Section 2 describes the proposed framework. The experiments and results are discussed in Section 3. Finally, Section 4 concludes the paper.

2. Proposed methodology

The proposed methodology consists of two phases, that is, a dictionary training phase and an SR phase. According to the problem posed in Eq. (3), we can generate an HR output A from a given LR image B by applying the recovered sparse coefficients X with an overcomplete dictionary G. The pictorial demonstration of the proposed framework is shown in Figure 1. The details about dictionary learning and SR are described in the following subsections.

2.1. Dictionary training

Effective image representation plays a vital role in image SR reconstruction. Various dictionary learning algorithms have been developed to replicate the image representations from the observed training images. From the last decade, K-SVD dictionary learning procedures have been deployed successfully for image de-noising and super resolution [11]. Literature reviews indicate that an appropriate modification both in sparse recovery and in dictionary updating stages can improve the performance of the dictionary learned through the K-SVD dictionary training procedure [9, 15].

The training procedure of the K-SVD consists of two parts: 1) sparse recovery method and 2) dictionary updating. The objective function for sparse recovery posed in Eq. (3) can be decoupled into N distinct problems of the form

$$\min_{x_i} \left\| \boldsymbol{y} - Gx_i \right\|_F^2 \; : \; \sum_i \left\| x_i \right\|_0 \leq \mathcal{K}$$
$$\text{for } i = 1,2,...,N \tag{4}$$

where $y \in R^n$ is the training patch randomly taken from the HR training image Y. The image patch y is involved in dictionary learning if the standard deviation σ of its intensity is greater than a threshold Δ, i.e., $\sigma > \Delta$. The L_0-minization in this problem is solved by applying ROMP [13]. ROMP is a simple regularized extension of OMP which has the transparency and greediness of OMP and the robustness of the L_1-minimization. In ROMP, the local approximation property recommends using the K biggest coordinates of the observation vector u=G*x instead of one biggest coordinate as OMP does. The observation vector u does better as a local approximation while reconstructing the HR patch a from the measurement x =Ga. ROMP emphasizes that the selected coordinates I must be more regular. Thus,
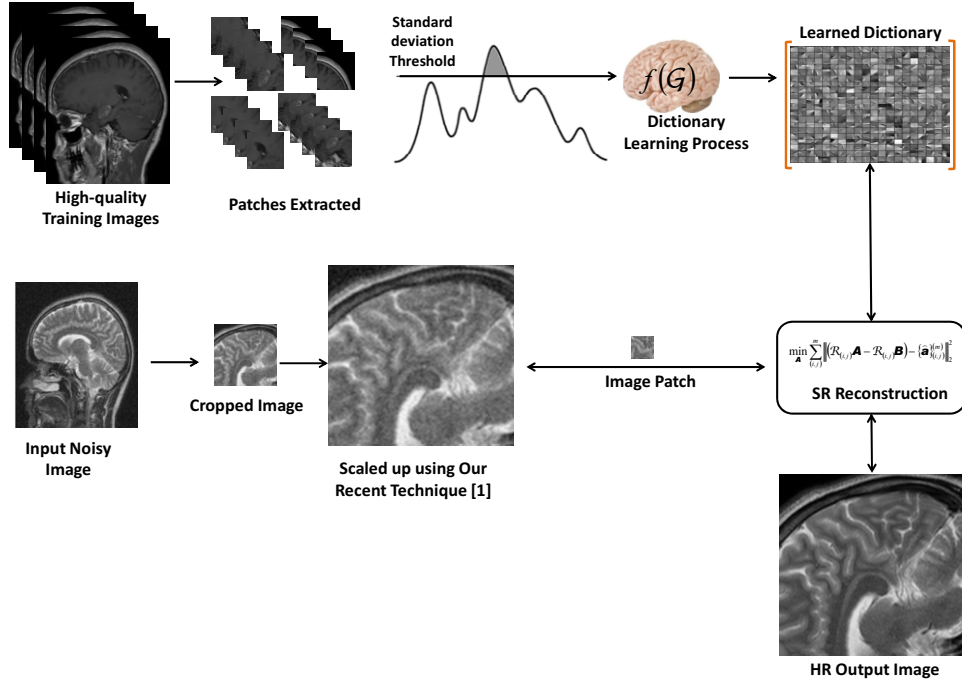
Fig. 1. Framework of the proposed super-resolution reconstruction technique.

coordinates J_0 with maximum energy and comparable sizes are selected and are added with the index set: $I \leftarrow I \cup J_0$. For selected coordinates I, we solve a least square problem and update the residual r to eliminate any contribution of the coordinates in I:

$$x = \min_{z \in \Re^I} \|\mathbf{y} - Gz\|_2 \tag{5}$$

$$r = x - Gx \tag{6}$$

where residual r is initialized to y at the start. The regularization step does not involve combinatorial complexity and can be performed in linear time. For a detailed study of ROMP, the readers are referred to [13]. The algorithm for ROMP can be summarized in Algorithm 1.

Algorithm 1

Task: Generate a vector of the sparse representation coefficients.

Input: G, y, sparsity level K.

Output: r, I, and x.

Initialization: Set $t = 1$, r = y, and I = 0.

while (stopping criterion not met**) do**

    (i). *Identify:* Chose a set J of the K biggest coordinates of the observation vector u=G*x.

    (ii). *Regularize:* divide the set J into J_0 where $\|u_i\| \leq 2 . \|u_j\|$ *forall* $i, j \in J_o$ with maximal energy $\|u|_{J_0}\|_2$.

    (iii). *Update:*
- Set $I_{(t)} = I_{(t-1)} \cup J_o$.
- Compute x by solving the problem posed in Eq. (5).
- Calculate the new residual r using Eq. (6).

    (iv). $t = t + 1$.

In dictionary updating stage, the procedure of simultaneous atoms optimization described in [15] is incorporated that allows to simultaneously update an arbitrary subset of atoms and the corresponding non-zero coefficients of the sparse representation. Let $I \subseteq G=\{g_1,\ldots, g_n\} \in R^n \times^m$ be the index set of the atoms to be updated. $G_{:,I}$ and $X_{:,I}$ indicates the sub-matrices indexed by I. The optimization problem for simultaneous updates can be written as

$$\inf_{G_{:,I}:G \in G} f_I(G)$$

(7)

and the corresponding objective function of $f_I(G)$ can be written as

$$f_I(G)= \inf_{X_{I,:}:X \in V(\Omega)} \left\| \mathbf{Y}_t - G_{:,I} X_{I,:} \right\|_F^2$$

(8)

where $V(\Omega) = \{X \in \mathcal{R}^{m \times n}: X_{i,j}=0 \forall (i,j) \notin \Omega\}$ and these two steps, sparse coefficients recovery and dictionary update are repeated until an overcomplete dictionary is produced. The resultant dictionary G contains image representations such as oriented edges, lines, and bars instead of raw patches prototypes.

2.2.   Super-resolution

Suppose, we have to reconstruct an HR output image A from a given low-resolution image B. First, B is interpolated to the size of A using our recent adaptive interpolation technique [1].

$$\mathbf{B}' = \mathbf{Q}_{(u)}\mathbf{A} + \varepsilon_{(u)}$$

(9)

where $Q_{(u)}$ is the interpolation operator for scaling up and $\varepsilon_{(d)}$ is the interpolation error or the blurring effect of the scaling process. Then, m patches $\{\mathbf{b}'\}^{(m)}$ each of size $\sqrt{n} \times \sqrt{n}$ are extracted from locations (i, j) using an extraction operator R. ROMP algorithm is applied as described in the previous section to seek the vectors of sparse representation coefficients $\{x\}_{(i,j)}^{(m)}$ for each patch in $\{\mathbf{b}'\}_{(i,j)}^{(m)}$. The corresponding HR patches $\{\hat{\mathbf{a}}\}_{(i,j)}^{(m)}$ for each patch in $\{\mathbf{b}'\}_{(i,j)}^{(m)}$ are approximated as

$$\{\hat{\mathbf{a}}\}_{(i,j)}^{(m)} = G\{x\}_{(i,j)}^{(m)}$$

(10)

The final HR image A can be reconstructed by solving the following optimization problem

$$\mathbf{A} = \min_{\mathbf{A}} \sum_{(i,j)}^{m} \left\| \left(\mathcal{R}_{(i,j)}\mathbf{A} - \mathcal{R}_{(i,j)}\mathbf{B}\right) - \{\hat{\mathbf{a}}\}_{(i,j)}^{(m)} \right\|_2^2$$

(11)

The problem posed in (12) iterates until, we get an HR image A of high quality.

3. Experimental results

In this section, we conducted a variety of experiments to prove the effectiveness of the proposed scheme practically. For this purpose, we acquired brain MR images of different views (i.e., coronal view, axial view, and sagittal view) from three datasets: 1) PIMS Hospital (PIMSH) [16], 2) Brain Web Tumor Repository (Brain-Web) [17], and 3) Harvard Medical School brain MRI dataset (Harvard-MRI) [18]. These datasets contain varieties of brain MR images of high quality more suitable for training and evaluation purposes. The dictionaries for the proposed SR reconstruction scheme and SR-KF were trained from 10^6 patches of size 9×9 sampled independently from the training datasets of brain MR images. Unless explicitly mentioned, the size of the dictionary was maintained as 2^{10} in all experiments. We modified the algorithm of SR-KF [11] according to the experiment as it was designed for key-frames extracted from video sequence captured via wireless sensor network.

Figure 2 shows the qualitative evaluation of the proposed scheme, compared with a standard bicubic scheme and with our previous SR scheme, SR-KF [11]. Both SR-KF and the proposed SR schemes have the capability of SR and denoising simultaneously. However, bicubic technique lacks this property of denoising and can only scale up the input image. The visual quality of the proposed scheme is much better than bicubic, but slightly better than SR-KF. This superiority of the proposed scheme was also quantitatively proved, computing peak signal to noise ratio (PSNR) and structure similarity index metric (SSIM) [19] as shown in Table 1.
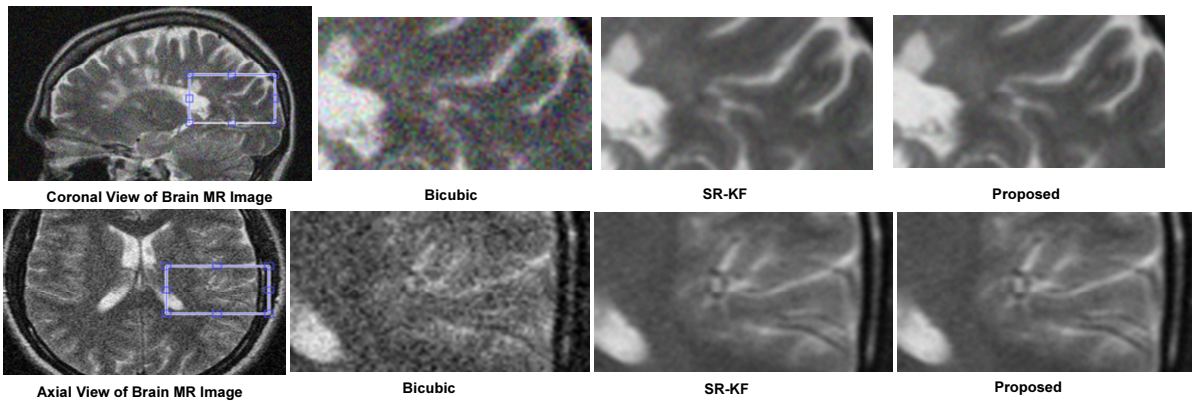

Fig. 2. Qualitative evaluation of the proposed technique on noisy (standard deviation Gaussian noise: $\mu = 0$, $\sigma = 0.01$) brain MR images. The cropped part of the distorted brain MR image is reconstructed $3\times$ using bicubic, SR-KF, and the proposed schemes.

Table 1
Quantitative evaluation: Accuracy supremacy of the proposed scheme over bicubic and SR-KF

| Brain MR images		1	2	3	4	5	6	7	8	9	10
Bicubic	PSNR	19.71	20.30	18.78	22.33	17.91	19.28	23.98	16.01	17.13	23.59
	SSIM	0.681	0.701	0.647	0.726	0.616	0.631	0.678	0.502	0.612	0.698
SR-KF	PSNR	25.46	26.50	24.02	27.53	23.50	24.39	28.98	25.77	26.50	28.94
	SSIM	0.829	0.837	0.792	0.857	0.706	0.730	0.750	0.706	0.760	0.852
Proposed	PSNR	28.34	29.11	27.90	30.12	26.39	26.82	31.18	28.28	29.14	31.94
	SSIM	0.881	0.894	0.864	0.911	0.796	0.851	0.868	0.781	0.865	0.958
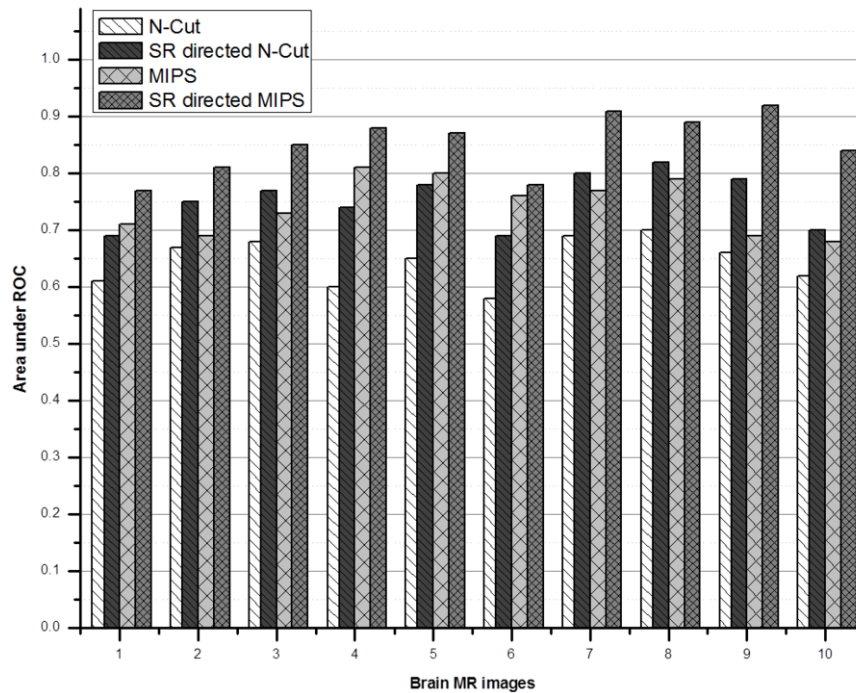
Fig. 3. Area under ROC curve (AUC) achieved by N-Cut and MIPS segmentation techniques. SR directed means that the segmentation technique was applied on the brain MR image super-resolved by the proposed SR technique.

Furthermore, the boosting effect of the proposed SR technique was evaluated in segmenting the brain tumor in MR images. The collected brain MR images were distorted by Gaussian noise, and two state-of-the-art segmentation schemes *normalized cut* (N-Cut) [20] and *medical image perception-based segmentation* (MIPS) [21] were taken into account to segment them. A standard parameter *area under ROC curves (AUC)* [21] was used for the objective assessment of the brain tumor segmentation results. Readers can find more detail about AUC in our recent article [21]. AUC computes a value in the range 0-1. A value approaching 1 indicates perfect segmentation and AUC=0 shows that the segmentation is completely incorrect, no overlap on tumor region (i.e., ground truth). AUC scores listed in Figure 3 depict the effect of the proposed SR technique as a pre-processing step in improving significantly the accuracy of two recent segmentation techniques. It can be easily concluded that the proposed SR technique makes the segmentation process more accurate.

4. Conclusion

This paper presented a novel scheme toward brain MR image SR reconstruction using the modified version of K-SVD trained dictionary. The sparse coding algorithm of the K-SVD is replaced by an efficient regularized OMP which has the transparency and greediness of OMP and the robustness of the L_1-minization. This two in one nature of ROMP enhances the learning process of the K-SVD dictionary in terms of features learning. In the SR reconstruction step of the proposed scheme, ROMP is used to determine the vector of sparse coefficients of the underlying patch. This vector of sparse coefficients is multiplied with trained dictionary to generate the HR output. The final HR image of

high-quality is generated through optimization process over learned dictionary. Both qualitative and quantitative analysis proved validity of the proposed SR reconstructed scheme.

In future work, we will investigate enhancing the capability of the dictionary to handle various types of distortion such as motion noise. The proposed algorithm will be elaborated for use in other applications such as video SR and de-noising, compressive sensing and representation and synthesis of visual texture. There are also additional improvements that can be made to the sparse coding Algorithm.

Acknowledgment

References

[1] M. Sajjad, N. Ejaz and S.W. Baik, Multi-kernel based adaptive interpolation for image super-resolution, Multimedia Tools and Applications **72** (2012), 2063-2085.

[2] M. Sajjad, N. Ejaz, I. Mehmood and S.W. Baik, Digital image super-resolution using adaptive interpolation based on Gaussian function, Multimedia Tools and Applications, 2013, 1–17. doi 10.1007/s11042-013-1570-1.

[3] A. Rueda, N. Malpica and E. Romero, Single-image super-resolution of brain MR images using overcomplete dictionaries, Medical image analysis **17** (2013), 113–132.

[4] J. Yang, J. Wright, T.S. Huang and Y. Ma, Image super-resolution via sparse representation, IEEE Transactions on Image Processing **19** (2010), 2861–2873.

[5] P.R. Gill, A. Wang and A. Molnar, The in-crowd algorithm for fast basis pursuit denoising, IEEE Transactions on Signal Processing **59** (2011), 4595–4605.

[6] W. Dong, D. Zhang, G. Shi and X. Wu, Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization, IEEE Transactions on Image Processing **20** (2011), 1838–1857.

[7] H. Lee, A. Battle, R. Raina and A.Y. Ng, Efficient sparse coding algorithms, Advances in Neural Information Processing Systems **19** (2007), 801:1–801:8.

[8] R. Zeyde, M. Elad and M. Protter, On single image scale-up using sparse-representations, Proceedings of 7th Conference on Curves and Surfaces, Avignon, France, June 2010, pp. 711–730.

[9] M. Aharon, M. Elad and A. Bruckstein, K-SVD: An algorithm for designing of over-complete dictionaries for sparse representation, IEEE Transactions on Signal Processing **54** (2006), 4311-4322.

[10] S. Yang, M. Wang, Y. Chen and Y. Sun, Single-image super-resolution reconstruction via learned geometric dictionaries and clustered sparse coding, IEEE Transactions on Image Processing **21** (2012), 4016–4028.

[11] M. Sajjad, I. Mehmood and S.W. Baik, Sparse representations-based super-resolution of key-frames extracted from frames-sequences generated by a visual sensor network, Sensors **14** (2014), 3652–3674.

[12] G. Rath and C. Guillemot, A complementary matching pursuit algorithm for sparse approximation, Proc. European Signal Process. Conf., Lausanne, Switzerland, 2008.

[13] D. Needell and R. Vershynin, Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit, Foundations of Computational Mathematics **9** (2009), 317–334.

[14] N.H. Nguyen and T.D. Tran, The stability of regularized orthogonal matching pursuit algorithm, 2007. http://www.dsp.ece.rice.edu/cs/ Stability-of-ROMP.pdf.

[15] W. Dai, T. Xu and W. Wang, Simultaneous codeword optimization (SimCO) for dictionary update and learning, IEEE Transactions on Signal Processing **60** (2012), 6340–6353.

[16] P.I.o.M. Sciences, http://www.pims.gov.pk/radiology.htm, Pakistan, 2010, Radiology Department, Department of Radiology, Pakistan Institute of Medical Sciences. http://www.pims.gov.pk/radiology.htm

[17] Brainweb, BrainWeb: Simulated Brain Database, Online: http://brainweb.bic.mni.mcgill.ca/cgi/brainweb2

[18] M.D. Keith, A. Johnson and J. Alex Becker, The Whole Brain Atlas, Harvard Medical School, Brain MRI Database. http://www.med.harvard.edu/aanlib/

[19]Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Transactions on Image Processing **13** (2004), 600−612.

[20]J. Shi, and J. Malik, Normalized cuts and image segmentation. Pattern Analysis and Machine Intelligence, IEEE Transactions on 22 (2000), 888−905.

[21] I. Mehmood, N. Ejaz, M. Sajjad and S.W. Baik, Prioritization of brain MRI volumes using medical image perception model and tumor region segmentation, Computers in Biology and Medicine **43** (2013), 1471−1483.